International Journal of Scientific Research in Science and Technology
Print ISSN: 2395-6011 | Online ISSN: 2395-602X (www.ijsrst.com)
doi : https://doi.org/10.32628/IJSRST

Detection of Fast Moving Vehicles In Aerial Videos Using Machine Learning Techniques

Dr. J Madhavan*[1] , Dr. A.Karthikayen[2], Dr. M. Venkatesan[3], Dr. Bhaludra R Nadh Singh[4] , Dr. Bremiga Gopalan[5]
*1 Professor, Department of ECE, Bhoj Reddy Engineering College for Women, Vinay Nagar, Hyderabad, Telangana, India.
*2 Professor, Department of ECE, Visvodaya Engineering College, Kavali, SPSR Nellore,Andhra Pradesh, India
Email id: akarthi_mathi@yahoo.co.in.
*3 Professor, Department of ECE, Visvodaya Engineering College, Kavali, SPSR Nellore,Andhra Pradesh, India
Email id: hellomvenkat@gmail.com.
*4 Professor, Department of CSE, AVN Institute of Engineering and Technology, Hyderabad,Telangana, India
*5 Assistant Professor, Department of ECE, Bhoj Reddy Engineering College for Women, Vinay Nagar, Hyderabad, Telangana, India

Article Info
Volume 7, Issue 6
Page Number : 532-539

Publication Issue
November-December-2020

Article History
Accepted : 10 Nov 2020
Published : 10 Dec 2020

ABSTRACT

Intelligent Transportation Systems (ITS) permit us to have high exceptional site visitors facts to lessen the danger of probably vital situations. Conventional image primarily based site visitors detection strategies have difficulties obtaining precise photographs due to attitude and historical past noise, negative lights and weather conditions. In this paper, we suggest a brand new method to correctly phase and track cars. After disposing of perspective the use of Modified Inverse Perspective Mapping (MIPM), Hough remodel is carried out to extract avenue lines and lanes. Then, Gaussian Mixture Models (GMM) are used to segment transferring itemsand to tackle vehicle shadow results, we observe a chromacity-based totally strategy. Finally, performance is evaluated via three one of a kind video benchmarks: personal recorded videos in Madrid and Tehran (with distinct weather situations at urban and interurban regions); and two famous public datasets (KITTI and DETRAC). Our effects imply that the proposed algorithms arestrong, and greater accurate in comparison to others, mainly when going through occlusions, lighting fixtures versions and climate situations.

Keywords : Historial, Camera Motion, machine learning techniques.

I. INTRODUCTION

Aerial videos acquired by airplanes or fixed-wing unmanned aerial vehicles (UAVs) are a good source of data for ground surveillance. Such kind of data can be used for applications, such as automatic traffic monitoring, border surveillance, or protection of critical infrastructures. In this article, we will focus on video data acquired with a frame rate of 25 Hz by a small UAV in top-down view at an altitude of 400 m. The ground coverage is up to 0.5 km2 with a ground sampling distance (GSD) of about 0.3 m/pixel. Such parameters are usual for so-called full-motion video (FMV) sequences. Additionally, there is one specific property that we can find in top-view videos with small-sized objects on the ground that are observed; there is no significant change in appearance of the objects in the image over time. This is crucial for the successful application of image stacking. Image stacking is a well-known method that is used to improve the quality of images in video data. A set of consecutive images is aligned by applying image registration and warping.1 In the resulting image stack, each pixel position contains redundant information about its intensity or color value. This redundant information can be used to perform image fusion to suppress image noise,2,3 handle motion blur,4–6 detect independent motion in images using background models,7,8 or even enhance the spatial image resolution as done in multi frame super-resolution.9–12 However, the performance of the above-mentioned methods is usually dependent onthe assumption that the scene is static and that the viewing angle does not change significantly over time. At the same time, moving objects can get blurred or distorted as the images inside the image stack are registered and aligned with respect to (w.r.t.) the stationary background. Hence, moving objects need to be handled explicitly as this is done for super-resolution.13,14 In this article, we apply image stacking in an innovative way: image alignment is not performed w.r.t. the stationary background but to small moving objects instead. In our aerial FMV data, moving objects are mainly cars but also other vehicles on the ground such as motorcycles. Then, we apply a pixelwise temporal filter (e.g., a median filter as done in Ref. 3) to the image stack to remove noise and compression artifactsfrom the moving objects. More interestingly, we intentionally add motion blur to the background that surrounds the moving objects. This effect is visualized in Fig. 1. We aim at detecting the moving vehicle that is marked with a red cross. As we can see in thezoomed original image (left box inside each image), there are potential distractors, such as parked and overtaking vehicles. By applying our proposed image stacking method (right box inside each image), it takes only 28 images (~1 s) to blur the distractors effectively. Although camera and objects are moving, the observed vehicle's appearance is nearly constant over time due to the top-view camera angle and dueto the large distance between camera and scene. In this way, stationary image content is removed that can modify the observed object's appearance and, thus, can disturb the detection and segmentation process, such as (1) partial occlusions by trees, power supply lines, or buildings, (2) stationary objects close to the observed object, such as parked vehicles or buildings, and (3) street textures, such as cobblestonesor road markings. To demonstrate the effectiveness of the proposed approach, two baseline object detection and segmentation algorithms are implemented, improved by our image stacking approach, andevaluated using FMV sequences that contain occlusions, parked vehicles, and street textures.

The main contributions of this article are:

1. the introduction of image stacking for moving ground objects observed by a moving airborne camera to intentionally blur the surrounding stationary background,
2. handling and management of multiple potentially overlapping image stacks that occur when observing multiple moving ground objects in parallel, and
3. how to use the proposed image stacking approach to improve standard moving object detection and segmentation methods.

We briefly mentioned the proposed image stacking approach in prior work already,15 but no details were given, such as how image stacks can be initialized, how multiple image stacks can be managed, and how our proposed approaches can be implemented efficiently. Furthermore, we provide an extensive study of the image stacking parameters thatcontribute most to the object detection and segmentation performance.

II. RELATED WORK

[1] Multiple Object Tracking Performance: The CLEAR MOT Metrics

Metrics that describe, he quality and key characteristics in numerous object tracking systems must be studied and compared in accordance to carefully analyse and evaluate their performance. Regrettably, there has yet to be agreement on such a range of generally valid measures. They present two new measures for evaluating MOT systems in this paper. Multiple object tracking precision (MOTP) as well as multiple object tracking accuracy (MOTA) are suggested benchmarks that can be used for a variety of monitoring activities & permit for objective contrast of tracking systems' primary features, likeaccuracy at locating targets, precision at recognising target configurations, but also way to detect targets on consistent bases. They put the proposed metrics to the test in a series of global evaluation workshops to see how useful and expressive they were. The CLEAR workshops in 2006 and 2007 featured a wide range of monitoring activities whereby a big number ofmodels were tested & evaluated. Their studies findings reveal that its suggested measures accurately reflect the numerous methods' qualities and shortcomings in a simple and direct manner, helps in easy evaluation in performance, thus relevant toward a wide range of circumstances.

[2] Fully-Convolutional Siamese Networks for Object Tracking

Traditionally, the problem of arbitrary target tracking was tackled by developing a system of the targets arrival entirely online, with only the video as training data. Despite their effectiveness, these approaches' online-only methodology limits using depth information which can be studied. Many efforts have actually been developed towards harnessing deep convolutional networks' descriptive ability. Once the target to monitor isn't determined ahead of time, Stochastic Gradient Descent online is required in adjusting the network's parameters, risking overall system's speed. For object detection in video, a basic tracking method is combined with a novel fully-convolutional Siamese network that has been trained end-to-end on the ILSVRC15 dataset. The tracker reaches state of art success in various tests with the minimal brevity. It works at fps that are faster than actual.

[3] Simple Online And Realtime Tracking

This research looks at a realistic approach for monitoring many items, with the primary objective of associating objects successfully for online and real-time operations. The study claims that recognition ability is a critical component in determining detection accuracy, with modifying its detector boosting tracking efficiency by up as 18.9 percentage. In contrast to many batch-based tracking systems, this research focuses on online tracking, where the tracker is only shown detections from the previous and current frames. Despite just employing a simple mix of existing techniques such as the Kalman Filter and the Hungarian algorithm for the tracking components, this approach achieves tracking accuracysimilar to state-of-the-art online trackers. This research looks at a realistic approach for monitoring many items, with the primary objective of associating objects successfully for online and real-time applications. The study claims that detection quality is a critical factor in determining tracking performance, with changing the detector boosting tracking performance by up to 18.9%. The trackeralso updates at a rate of 260 Hz, which is nearly 20 times faster than other state-of-the-art trackers due tothe simplicity of the tracking method.

[4] High-Speed Tracking-by-Detection without Using Image Information

As the performance in object detectors increases, the foundation for a tracker becomes significantly more trustworthy. The problems for a successful trackerhave changed as a result of this, as well as the increased use of higher frame rates. As a result of this shift, considerably simpler tracking algorithms may now compete with more complex systems for a fraction of the processing cost. This paper outlines and illustrates such a method by conducting extensivetests with a range of object detectors. The proposed technique can easily operate at 100K fps on the DETRAC vehicle tracking dataset, beating the state-of-the-art. The notion of a passive detection filter is used to analyse a very simple tracking technique in this research. Due to its modest computing footprint, the suggested approach can serve as a basic predictive model for other trackers and provide an appraisal of the necessity of additional efforts in the tracking algorithm. It also permits reviewing tracking benchmarks to evaluate if the specific concerns they indicate (for instance, missed detections, frame rate, etc.) are within the capabilities of existing algorithms.

[5] Cascade R-CNN: Delving into High Quality Object Detection

The this study, they presented the Cascade R-CNN, a multi-stage object recognition framework for developing high-quality object detectors. Overfitting during learning and quality disparity during inference have both been demonstrated to be avoided using this design. On the hard COCO and famous PASCAL VOC datasets, the Cascade R-substantial CNN's and consistent recognition improvements show that effective object detection necessitates modelling and knowledge of several corroborating aspects. The Cascade RCNN has been shown to work with a wide range of object detection architectures. They hope it will be useful in a variety of future object detection research projects.

[6] Real-Time 'Actor-Critic' Tracking

Object tracking is one of the most important step in object detection and tracking. In this paper Actor-Critic framework been used, where the 'Actor' model seeks to infer the best option in a continuous action space, causing the tracker to move the bounding box to the object's current location. The 'Critic' model is used for offline training to create 'Actor-Critic' framework along with reinforcement learning as well as a Q-value for directing both the 'Actor' and 'Critic' deep network learning processes. Visual tracking is viewed as a dynamic search process in which the 'Actor' model outputs only one action to locate the tracked object in each frame. Offline training of better policy for finding the best result is done using reinforcement learning. Furthermore, the 'Critic' network serves as a verification system for both offline and online instruction. Using popular benchmarks, The suggested tracker gets contrasted to certain state of art trackers, as well as the stimulation finding reveal that it performs well in actual.

[7] Hybrid Task Cascade for Instance Segmentation At instance level, segmentation process is basic

computer vision job which identifies objects per-pixel. In real-world settings like automated driving and video surveillance, precise and reliable feature extraction is challenging to achieve. Cascade wouldbe a basic but efficient design which has increased results on a wide range of workloads. A basic Cascade R-CNN and Mask R-CNN combination produces just a little boost. The secret to good instance segmentation cascade would be to completely use inverse interaction across detection as well as segmentation so that to discover a more effective technique.

They propose Hybrid Task Cascade (HTC) in this paper, that is different in two key ways: (i) it intertwines these two tasks for simultaneous multi-stage computation, rather than conducting cascaded refinement on them individually; and (ii) employs a convolutional section in order to give spatial features, that help distinguish difficult frames in cluttered background. Bounding box analysis plus masked predictions is coupled inside a multi-tasking way at each step of HTC. At certain stages, easily applicable within its masked sections are also given - the masked characteristics in every step are combined and supplied to the next. The whole design increases data flow inside the activities as well as stages, resulting in good refined predictions at all levels and more reliable forecasts overall. HTC is simple in setting up & may be programmed from beginning to end. It gained 2.6 % & 1.4 % greater masked AP than that of the Masked R-CNN as well as Cascade Masked R-CNN benchmarks, respectively, on difficult COCO benchmark.

III. III.PROPOSED SYSTEM

In this section, we describe our proposed method in detail. As discussed below, we divide our developed method into the following interconnected steps along with a brief description. Figure 2 shows the flow of the proposed method. In addition, Algorithm 1 shows more details of our developed method. To test our method, we gather our own dataset from challenging Pakistani traffic environments. This dataset was collected over a period of two months in different cities of Pakistan. As shown in Figure 2 that the gathered data is preprocessed and augmented. Later it is trained by our model. Meanwhile, the YOLO-v5 model is built and trained. Our collected data is from an unknown distribution in Pakistani traffic. Therefore, it is now tested on the YOLO-v5 model. After the YOLO-v5 is applied, we then investigateand analyse our detector. To aid readers' understanding, below we describe the steps and details of our developed method.

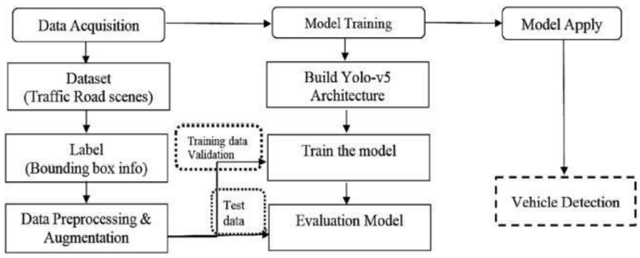

Fig 1: Proposed Architecture

To begin with the proposed algorithm, we initially acquire data. First we deal with different conditions on highways. For example, we come across the multi-class objects, such as different types of vehicles, motor bikes, and pedestrians on the roads. Similarly, we also faced severe and crucial challenges, such as massive traffic jams and overlapped vehicles. Therefore, to systematically acquire the data as shown in line (6) of Algorithm 1, we collected the dataset under two different situations, which are (i) High Density Traffic (HDT) scenes that contains multiple objects in an image and (ii) the Low-Density Traffic (LDT) scene that contains only one class per image, with zero overlaps. For improved training, the images of theLDT and the HDT dataset are placed separately. The LDT Scenes: This dataset was gathered from daily real-time traffic places, for example open parking lots, less crowded roads, and places with fewer crowds. The objective of assembling this dataset is to separately train the model on each class. We collecteda total of 600 images from three classes, which are cars, motor cycles, and pedestrians.

IV. EXPERIMENTAL ANALYSIS

This section presents the detailed simulation results. Extensive experiments are carried out on Google Colaboratory (Colab) platform. The Google Colab provides Intel Xeon CPU with a clock speed of 2.3 GHz and up to 16 GB of RAM. Moreover, the Google Colab also provides NVIDIA K80 or T4 GPU. We use Python V3.6 as a simulation tool for different vehicle datasets as described in subsequent sections. To investigate the performance of vehicle detection methods on different datasets, we select 14 state-of- the-art vehicle detector evaluations and comparisons with the proposed method in terms of accuracy and execution time. All of the compared approaches have been trained on the same training data from each of the PKU, COCO, and DAWN datasets.

V. CONCLUSION

In this research, we have proposed a robust method for extracting real information from traffic cameras. The research focused on different issues, namely removing perspective, automatic locating of lines and lanes, vehicle detection and extracting features of the vehicles. As the main contribution of this research, we have proposed a method to remove perspective without any harmful effect on the real information The sliding window technique is a properly desirable technique for vehicle detection in aerial movies. In our experiments, we display that it could outperform detection algorithms based totally on item segmentation especially in urban scenes with many vehicles riding on busy streets. Parameters of the sliding window method that contribute maximum to the detection and processing performance are recognized and optimized: we endorse (1) to use ChnFtrs + AdaBoost as automobile version, (2) to rescale the photograph with best three distinct scales and best in width route with fixed top, and (3) to optimize the runtime with gentle cascades, subsampling, and reducing the wide variety of susceptible classifiers in the AdaBoost version. In this way, we achieve detection fees of 88 % with simplest 2 % of FP detections throughout exceptional datasets and a mean processing time less than 40 ms according to body on trendy hardware in scenes with up to twenty transferring motors. The low FP price together with detection confidences furnished by using the classifier make sliding window based object detection suitable for a combination with a couple of item monitoring processes that depend on initial detections.

VI. REFERENCES

1. R. Benenson, M. Mathias, R. Timofte, and L. Van Gool. Pedestrian detection at 100 frames per second. In Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), 2012. 2, 3, 4,6
2. X. Cao, C. Wu, J. Lan, P. Yan, and X. Li. Vehicle Detection and Motion Analysis in Low-Altitude Airborne Video Under Urban Environment. IEEE Transactions on Circuits and Systems for Video Technology, 21(10):1522–1533, Oct. 2011. 1, 2, 6
3. H.-Y. Cheng, C.-C.Weng, and Y.-Y. Chen. Vehicle Detection in Aerial Surveillance Using Dynamic Bayesian Networks. IEEETransactions on Image Processing, 21(4):2152– 2159, Apr. 2012. 8
4. M.-M. Cheng, Z. Zhang,W.-Y. Lin, and P. H. S. Torr. BING: Binarized Normed Gradients for Objectness Estimation at 300fps. In Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), 2014. 2
5. N. Dalal and B. Triggs. Histograms of Oriented Gradients for Human Detection. In Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), pages 886–893, 2005. 1, 2, 4, 6
6. S. J. Davey, M. G. Rutten, and B. Cheung. A Comparison of Detection Performance for Several Track-Before-Detect Algorithms. In Proceedings of the 11th International Conference on Information Fusion (FUSION), 2008. 3
7. P. Doll´ar, S. Belongie, and P. Perona. The Fastest Pedestrian Detector in the West. In Proceedings of the British Machine Vision Conference (BMVC), 2010. 2, 4
8. P. Doll´ar, Z. Tu, P. Perona, and S. Belongie. Integral Channel Features. In Proceedings ofthe British Machine Vision Conference (BMVC), 2009. 2, 3, 6
9. P. F. Felzenszwalb, R. B. Girshick, D.McAllester, and D. Ramanan. Object Detection with Discriminatively Trained Part Based Models. IEEE Transactions on Pattern Analysis and Machine Intelligence, 32(9):1627–1645, Sept. 2010. 2
10. U. Franke, C. Rabe, H. Badino, and S. Gehrig. 6D-Vision: Fusion of Stereo and Motion for Robust Environment Perception. In W. G. Kropatsch, R. Sablatnig, and A. Hanbury, editors, Pattern Recognition, volume 3663 of Lecture Notes in Computer Science (LNCS), pages 216–223. Springer Berlin / Heidelberg, Sept. 2005. 3

11. Y. Freund and R. E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. Journal of Computer and System Sciences, 55(1):119–139, Aug. 1997. 4
12. A. Gaszczak, T. P. Breckon, and J. Han. Real-time people and vehicle detection from UAV imagery. In Proceedings of SPIE Vol. 7878, Intelligent Robots and Computer Vision XXVIII: Algorithms and Techniques, 2011. 2, 6
13. G. Gualdi, A. Prati, and R. Cucchiara. Multistage Particle Windows for Fast and Accurate Object Detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 34(8):1589–1604, Aug. 2012. 2
14. C. Harris and M. Stephens. A combined corner and edge detector. In In Proceedings of the Fourth Alvey Vision Conference, 1988. 3
15. R. Hartley and A. Zisserman. Multiple-View Geometry in Computer Vision. Cambridge University Press, 2 edition, Mar. 2004. 3
16. C. Heng, S. Yokomitsu, Y. Matsumoto, and H. Tamura. Shrink boost for selecting multi-LBP histogram features in object detection. In Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), 2012. 6
17. N. Japkowicz and M. Shah. Evaluating Learning Algorithms: A Classification Perspective. Cambridge University Press, 2011. 5
18. R. Kasturi, D. Goldgof, P. Soundararajan, V. Manohar, J. Garofolo, M. Boonstra, V. Korzhova, and J. Zhang. Framework for Performance Evaluation of Face, Text, and Vehicle Detection and Tracking in Video: Data, Metrics, and Protocol. IEEE Transactions on Pattern Analysis and Machine Intelligence, 31(2):319–336, 2009. 5
19. B. Leibe, A. Leonardis, and B. Schiele. Robust object detection with interleaved categorization and segmentation. International Journal of Computer Vision, 77:259–289, May 2008. 2
20. T. T. Nguyen, H. Grabner, H. Bischof, and B. Gruber. Online Boosting for Car Detection from Aerial Images. In Proceedings of the IEEE International Conference on Research, Innovation and Vision for the Future (RIVF), 2007. 1, 2, 6
21. C. Papageorgiou and T. Poggio. A Trainable System for Object Detection. International Journal of Computer Vision (IJCV), 38(1):15– 33, June 2000. 1, 2
22. V. Reilly, H. Idrees, and M. Shah. Detection and Tracking of Large Number of Targets inWide Area Surveillance. In Proceedings of the European Conference on Computer Vision (ECCV), 2010. 1
23. I. Saleemi and M. Shah. Multiframe Many-Many Point Correspondence for Vehicle Tracking in High Density Wide Area Aerial Videos. International Journal of Computer Vision (IJCV), 104(2):198–219, Sept. 2013. 1
24. J. Shi and C. Tomasi. Good features to track. In Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), 1994. 1, 3
25. M. Siam and M. ElHelw. Robust Autonomous Visual Detection and Tracking of Moving Targets in UAV Imagery. In Proceedings of the IEEE International Conference on Signal Processing (ICSP), 2012. 1
26. M. Teutsch, W. Kr¨uger, and J. Beyerer. Evaluation of Object Segmentation to Improve Moving Vehicle Detection in Aerial Videos. In Proceedings of the IEEE International Conference on Advanced Video and Signal-Based Surveillance (AVSS), 2014. 1, 2, 3, 6
27. S. T¨urmer, F. Kurz, P. Reinartz, and U. Stilla. Airborne Vehicle Detection in Dense Urban Areas Using HoG Features and Disparity Maps. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (JSTARS), 6(6):2327–2337, Dec. 2013. 1, 2, 6

28. P. Viola and M. Jones. Robust Real-time Face Detection. International Journal of Computer Vision, 57(2):137–154, 2004. 1, 2
29. Y. Wei and L. Tao. Efficient Histogram-Based Sliding Window. In Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), 2010.2
30. J. Xiao, H. Cheng, H. Sawhney, and F. Han. Vehicle Detection and Tracking in Wide Field-of-View Aerial Video. In Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), 2010. 1

Cite this Article

Dr. J Madhavan, Dr. A.Karthikayen, Dr. M. Venkatesan, Dr. Bhaludra R Nadh Singh , Dr. Bremiga Gopalan, "Detection of Fast Moving Vehicles In Aerial Videos Using Machine Learning Techniques", International Journal of Scientific Research in Science and Technology (IJSRST), Online ISSN : 2395-602X, Print ISSN : 2395-6011, Volume 7 Issue 6, pp. 532-539, November-December 2020.

Journal URL : https://ijsrst.com/IJSRST2183313